Hydrol. Earth Syst. Sci. Discuss., 7, 6525–6551, 2010
www.hydrol-earth-syst-sci-discuss.net/7/6525/2010/
doi:10.5194/hessd-7-6525-2010

Hydrology and Earth System Sciences Discussions

Modeling moisture fluxes using artificial neural networks: can information extraction overcome data loss?

A. L. Neal[1], **H. V. Gupta**[1], **S. A. Kurc**[2], **and P. D. Brooks**[1]

[1]Department of Hydrology and Water Resources, University of Arizona, 1133 James E. Rogers Way, Tucson, AZ 85721, USA
[2]School of Natural Resources and the Environment, University of Arizona, 325 Biosciences East, P.O. Box 21004, Tucson, AZ 85721, USA

Received: 27 July 2010 – Accepted: 20 August 2010 – Published: 1 September 2010

Correspondence to: A. L. Neal (aneal@email.arizona.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Eddy covariance sites can experience data losses as high as 30 to 45% on an annual basis. Artificial neural networks (ANNs) have been identified as powerful tools for gap filling, but their performance depends on the representativeness of data used to train the model. In this paper, we develop a normalization method, which has similar performance compared to conventional training approaches, but exhibits differences in the timing of fluxes, indicating different and previously unused information in the data record. Specifically, the differences between half-hourly model fluxes, especially during summer months, indicate that the structure of the information content in the data changes seasonally, diurnally and with the rate of data loss. This variation between gap-filling models complicates the application of their output as consistent data sets for land surface modeling, and points to the need for improved data and models to address flux behavior at critical times. We advise several approaches to address these concerns, including use of separate models for day and nighttime processes and the use of multiple data streams at dawn, when eddy covariance may be particularly ineffective due to the timing of the onset of turbulent mixing.

1 Introduction

Automated field data collection often produces discontinuous data sets as a result of instrument malfunction, power failure, or various other technical and non-technical problems. These discontinuities prove especially problematic for micrometeorological measurements. The expanded use of micrometeorological systems for ecological studies (Baldocchi et al., 2001) has resulted in an increased interest in methods to interpolate values for missing data. Eddy covariance measures landscape-scale energy and mass fluxes in a wide variety of ecosystems at high temporal resolution (generally 30 min accumulations). Towers are employed in a wide array of geographic regions including agricultural lands, temperate forests, tropical rainforests, and a range of arid and semi-arid landscapes (Baldocchi et al., 2001; Kurc and Small, 2007; Scott et al., 2006; Wohlfahrt et al., 2008), leading to important insights into the nature of the soil-vegetation-atmosphere system.

Data acquired using eddy covariance typically has significant gaps caused by insufficient turbulent mixing (Blanken et al., 1998; Goulden et al., 1996) or the sensitivity or failure of equipment, as well as poorly identified source areas (Brown-Mitic et al., 2007). Such factors can lead to violations of the assumptions of the eddy covariance technique, resulting in data being discarded during processing. These gaps are often serially correlated to particular events or periods important for observation, such as extreme weather events or nighttime carbon exchange (Falge et al., 2001) and transpiration (Dawson et al., 2007; Fisher et al., 2007).

To develop daily, seasonal, and annual estimates of fluxes, a method to fill gaps by approximating values for missing data is crucial. Gap-filled flux data is also applied in land-surface modeling studies, which often require continuous data streams for parameter identification. The method used for gap filling should result in flux time series that provide realistic assessments of moisture flux, energy and carbon exchange. For example, carbon flux gap filling has been shown to alter estimates of annual carbon exchange, including changes in the source/sink behavior of the carbon flux (Moffat et al., 2007), which presumably better reflects the true behavior of the ecosystem.

Gap filling requires awareness of the nature of missing data; i.e. the user should know what periods have missing data, why those data are missing, and what relationships will produce an appropriate estimate for that data. Gap-filling efforts should concentrate on times where flux records are critical, e.g. periods when flux data are important for modeling physical processes, such as the breakup of the stable nocturnal boundary layer, or when values may alter annual flux budgets, such as nocturnal fluxes. In all gap-filling scenarios, the presence of some valid data during similar periods is necessary to identify and model the appropriate relationships to replicate expected values when data are lost. Of course, at very high data loss rates, even a well-trained model may not be sufficient to reproduce the true behavior of a system.

HESSD
7, 6525–6551, 2010

Gap filling requires the use of a model to estimate missing data from existing reference data – numerous models for gap filling have been explored (Falge et al., 2001; Gove and Hollinger, 2006; Hui et al., 2004; Knorr and Kattge, 2005; Papale and Valentini, 2003). Analysis of several models by Moffat et al. (2007) suggests that artificial neural network (ANN) methods can provide substantial benefits for gap-filling studies; they perform similarly to other methods without requiring prior assumptions regarding model structure and are computationally less expensive. Moffat et al. (2007) also note in their conclusion the need to expand testing of gap filling models to a variety of different ecosystems, including arid sites. While arid sites may behave differently under physically based models due to differences in energy and water partitioning, ANNs should demonstrate similar model performance levels in humid and arid regions.

Neural network methods identify input-output relationships in a manner dependent on the information contained in the input and output data sets (MacKay, 2003). When training an ANN, using data from periods that can maximize information regarding the input-output relationship is key. In the case of gap filling flux records, this requires capturing the input-output pattern during low turbulent exchange, since these are often periods which have missing or filtered data. Valid data from periods of low turbulence, that is just above the filter threshold, are particularly valuable as a result (validity often established by a criterion such as friction velocity; Blanken et al., 1998; Goulden et al., 1996). By extracting the maximum available information from data in conditions near the filter threshold, we can improve the results of gap filling. This study considers an approach to treating data that modifies the distribution of data to extract information from data near the filter threshold.

In brief, gaps in flux records pose problems for the development of seasonal and annual estimates of evapotranspiration at the landscape scale. These gaps also make it difficult to conduct model-based investigations of forcing-response relationships at the land surface. Here, we investigate the use of an ANN framework to fill gaps in energy fluxes, with a particular focus on the probability distributions of flux data and the associated information applied during ANN model training. We apply a standardization

HESSD
7, 6525–6551, 2010
Modeling moisture fluxes using artificial neural networks
A. L. Neal et al.

technique that converts the probability density function of the flux data to an approximate normal distribution. This transformation moves extreme events toward the central tendency while expanding the region of near-zero fluxes where much of the sensitivity in gap-filled flux records is found (Falge et al., 2001). By altering the shape of the distribution, we can improve the ability of the ANN training algorithm to detect extreme events.

2 Methods

2.1 Data and model structure

Data used for this study were collected at a mesquite woodland site in Southeastern Arizona, near Tombstone, AZ (Scott et al., 2004). The data record spans three years (2001–2003) – the tower was not operational during the winter of 2001–2002. Data were filtered according to Scott et al. (2004). Time series of the latent heat flux (Fig. 1) contain data gaps typical of eddy covariance records, with both long (multiple-day to week) and short gaps (several hours to individual 30-min intervals). Similar patterns of missing data have been found in flux tower records from mesic and humid systems including a range of forest types (tropical and temperate, broadleaf and needleleaf; Falge et al., 2001; Moffat et al., 2007).

The ANN models used in this study compute estimates of latent and sensible heat fluxes based on inputs of precipitation, relative humidity, wind speed, air pressure, net radiation and temperature. Fluxes were filtered prior to modeling using a friction velocity (u^*) filter reported in (Scott et al., 2006). The input-output structure was similar to simple equations of evapotranspiration (e.g., Penman equation; Alavi et al., 2006) as well as land-surface models (Pitman, 2003). Input meteorological data come from a site less than 1 km from the tower and were gap-filled independently using a mean diurnal value method. After running the model, the latent heat fluxes were then converted into evapotranspiration (ET, mm/30-min) for ease of interpretation.

HESSD
7, 6525–6551, 2010

Model training used the mean absolute error (MAE) to account for the error structure of individual flux observations (Richardson et al., 2006). Training was performed on the data from 2003, comprising the longest near-continuous record in the data, and was validated over the entire three-year time series. Gap filling performance was evaluated on existing gaps in the data to assess the performance of models in a "real" gap filling application. Here, model performance on existing data is used to evaluate the predictive capability of the model.

2.2 ANN training and information content

2.2.1 Approaches to ANN training data

When using ANNs, pre-processing methods are used to transform input and output data onto the range ±1 (e.g., Matlab Neural Network Toolbox, MathWorks, Inc., Natick, MA). This ensures that the model will predict outputs based on the scaled variation in and among the data, rather than one dominant data stream of large magnitude. As a way to assess this scaling process, we examine the probability distribution function of the output training data (the measured values of ET) before applying the ANN for gap filling. Two different scaling techniques are applied, the first a simple scaling, reducing values to the range [−1 1] and the second making a standardized distribution, which is described below.

In order to standardize the flux data, the raw data is assumed to belong to a gamma distribution. The absolute value of the minimum is added to each point in the raw data so that the entire record is positive. After this shift, a gamma distribution was fit using a maximum likelihood method in the Matlab statistical functions (*gampdf* and *gamcdf*). These statistical functions were used to transform the data to a normal distribution using a method similar to the standardized precipitation index (SPI), after (Mishra and Desai, 2006). The normalization is achieved using an approximation for the normal

HESSD
7, 6525–6551, 2010

distribution; (see Mishra and Desai, 2006 for more details):

$$
\begin{aligned}
Z &= -\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \quad \text{for } 0 < H(x) \le 0.5 \\
Z &= \left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \quad \text{for } 0.5 < H(x) < 1
\end{aligned}
\tag{1}
$$

where, Z is the standardized index. The term, t, is calculated as:

$$
\begin{aligned}
t &= \sqrt{\ln\left[\frac{1}{(H(x))^2}\right]} \quad \text{for } 0 < H(x) \le 0.5 \\
t &= \sqrt{\ln\left[\frac{1}{(H(x))^2}\right]} \quad \text{for } 0.5 < H(x) < 1
\end{aligned}
\tag{2}
$$

where $H(x)$ is the cumulative gamma function, and the coefficients are:

$$
\begin{aligned}
&c_0 = 2.515517 \quad c_1 = 0.802853 \quad c_2 = 0.010308 \\
&d_1 = 1.432788 \quad d_2 = 0.189269 \quad d_3 = 0.001308
\end{aligned}
$$

The transformation alters the distribution of the latent heat flux data applied to the model as shown in Fig. 2. The results of this transformation are used directly as the training data for the model. The subsequent output is then converted back to values of latent heat flux using a third-order polynomial fit (r^2=0.98). By applying this transformation, we are able to alter the sampling patterns used by the ANN during training, and thus enhance the sampling from extreme values of latent heat flux.

2.2.2 Information theory and ANN training

To quantify the effects of different methods of data scaling, we consider the training data in terms of its information content, that is, the new behavior described to the model

HESSD
7, 6525–6551, 2010

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

by the data. Maximizing the information extracted from data should lead to the best possible model performance. We use the Shannon index to estimate the information associated with the frequency of certain values of latent heat flux. The Shannon index (Shannon, 1948), quantifies the amount of information, $h(x)$, characteristic of a value (or range), x, as a function of its probability $p(x)$, according to the expression:

$$h(x) = -p(x)\ln(p(x)) \quad (3)$$

and the global information content, H, for all values of x in the set, S:

$$H = \sum^{S} h(x) = -\sum^{S} p(x)\ln(p(x)) \quad (4)$$

As noted above, most implementations of ANNs, including the default settings of most software packages, simply rescale the distribution of the raw data to an interval [−1 1] (e.g. the *mapminmax* function in Matlab). Under this typical rescaling approach, the total Shannon index value of the distribution shown in Fig. 2b is 1.08. Values for each bin are shown in Table 1. To increase the information extracted from extreme low and high values of latent heat flux, the data were transformed from their original distribution into a (near) normal distribution, as described below. The resulting transformation yielded a total Shannon index of 1.77. At very low values of latent heat fluxes (less than 125 $\mathrm{W\,m^{-2}}$, or 0.1 mm/30 min), the information content in the original distribution is 1.05 compared to 1.37 in the standardized distribution.

We estimate the Shannon index, using Eqs. (3) and (4), in the training record. These index values describe the information content passed to the ANN during training. Equations (3) and (4) are not explicitly used to inform ANN training, but instead provide offline information about the characteristics of the data applied to the model.

2.3 Model performance

The two ANN models developed here were trained using 1) the typical approach of rescaling the latent heat flux data, referred to here as the "rescaled" model and 2) the

HESSD
7, 6525–6551, 2010

new approach which used a standardized approach, referred to as the "standardized" model. Both models were trained using MAE, as noted above. Model training used data from the year 2003, which represents the longest near-continuous subset of the data that incorporates a full range of seasonal behavior. A one-year data record is the shortest span which can reproduce seasonality without risk of overtraining at longer times (Neal, 2008).

After running the models, the output latent heat flux was converted to evapotranspiration (ET) for ease of comparison between models and to conventional ET measurements from other studies. Results are reported for the raw, unfilled data set (ET_{raw}), the rescaled model (ET_{rsc}) and the standardized model (ET_{std}).

Model performance was determined using several metrics: root mean squared error (RMSE), relative RMSE, MAE and Pearson's correlation. These metrics were calculated for the training period as well as the entire data record. Model residuals were also used to quantify model performance as a function of data loss. These residuals were calculated for each half hour interval and averaged for the entire record as well as seasonally (see below for a description of season delineation). The fraction of missing data for each half hour during these time periods was used to characterize data loss.

For comparison, uncertainty bounds were determined for the raw data during the training period. Because the tower was only operating during the growing season of 2001 and 2002, uncertainty estimates were only calculated based on the 2003 data. Uncertainty was determined for each 30-min interval using the difference between flux values on days of similar environmental conditions to identify the measurement uncertainty (Richardson et al., 2006).

As noted above, patterns in model performance were analyzed on daily and half-hourly time steps. Performance was also examined based on seasonal behavior. Seasons were defined as winter (December-January-February); monsoon, describing the North American Monsoon (Gochis et al., 2006) and identified by a climatologically defined rainy season (Kurc and Small, 2007) between late June and early September; and pre- and post-monsoon, the remainder of the year not contained in the other two seasons. These seasons are effectively winter, summer and spring/fall, but we use the climatologically defined monsoon to better identify moisture availability during the wet season.

3 Results

Model training resulted in similar performance for both models (Table 2). While the rescaled neural network model (ET_{rsc}) has a slightly lower RMSE and a slightly higher correlation, the difference in both is within the measurement uncertainty for the data. At a daily level, the difference in RMSE values for the two models is lower than the half-hourly RMSE, suggesting that the ANN model training may lead to compensation between under- and overestimation at different times during the day.

Time series of daily ET for the raw and model data during the year 2001 follow an expected seasonal trend (Fig. 3). The difference between the annual ET derived from the two neural network methods approximately 50 mm. Both of the model methods (ET_{rsc}, and ET_{std}) generally reproduce the seasonal pattern of ET at the daily level, though they differ in the magnitude of ET response to precipitation events, with ET_{std} producing more ET following storms but regressing to lower ET between storms.

Note that ET is greater than annual precipitation (253 mm in 2001). This difference is likely due to the influence of groundwater near the riparian corridor. Plant access to and use of groundwater at this site is discussed in several other studies, and has been linked both to inputs from the aquifer as well as bank storage following high flow events in the San Pedro River (Scott et al., 2006, 2004, 2008). The strong atmospheric demand on moisture from the semi-arid climate suggests that the availability of groundwater will have little impact on the ability to model ET based on meteorological variables. In terms of this study, omission of a groundwater term is not likely to affect either model in comparison to the other.

Comparing the two models in terms of daily ET (Fig. 4), the two models deviate from each other at daily ET values around 2 mm/day. These periods (region 2 in Fig. 4)

HESSD
7, 6525–6551, 2010

correspond to the pre- and post-monsoon season. At high values of daily ET, i.e. during the monsoon season, comparing the two models shows a high degree of scatter around the 1:1 line. At low values of ET, which generally occur during winter, the two models correspond well at the daily level.

Along with daily ET, the diurnal patterns of ET flux should be reproduced by a gap filling model. In the overall data set as well as in each of the three seasons, data loss is high during the nighttime and low, though still substantial around midday (Fig. 5). Rates of data loss generally follow the pattern of sunlight hours for each season. These rates of data loss, from the entire data set and each season independently, are used to estimate model performance as a function of data loss.

Applying the rate data loss as a predictor of model performance, using several forms of the model residual, the two models appear relatively consistent (Fig. 6). Residuals calculated include the absolute ($ET_{model}-ET_{raw}$) and relative (as a fraction of ET_{raw}) values as well as the residual normalized by the standard deviation of the data. The models both show slight improvement in the absolute value of residuals and worsening performance in terms of relative residuals. Most notably, however, the residuals from ET_{std}, which are much larger in magnitude than ET_{rsc} at low rates of data loss, actually indicate improved performance than ET_{rsc} through intermediate levels of data loss (between 40 and 70% loss, Fig. 6).

The diurnal pattern of model performance (Fig. 7) indicates poor model results during dawn in all three seasons and in the overall record. Errors during dawn and near-dawn periods are larger even than dusk and near-dusk times. Because the near-dawn period is associated with the breakup of stable nighttime air and the return of turbulent flux at the boundary layer, poor model performance is related to this change in the nature of the measured flux data.

HESSD
7, 6525–6551, 2010

4 Discussion and conclusions

At a qualitative level, the two models both reproduce the seasonal and event-based patterns in the data record. Timing of peak ET events occurs at similar times at the daily time-step when viewed on an annual level (Fig. 3). However, when examined more closely, the differences in model behavior become more noticeable and form the basis for the comparison in this study.

4.1 Model performance and information extraction

The standardized model, with a higher Shannon index value, should yield a trained model that better represent the observed pattern of ET, especially at low flux magnitudes. However, in terms of error metrics, the two models perform similarly (Table 2). The rescaled data appears to be biased toward lower values of ET, while the standardized model favors larger values of ET. In both cases, the performance of each model does not fully replicate the distribution of the raw data. However, the representativeness of any of the three distributions is complicated by the frequent loss of data, especially at low ET.

As shown here, ANNs as gap-filling tools are insensitive to data treatment (rescaled or standardized). The differences in performance between the two models (Table 2) are much smaller than the magnitude of the errors. Slight differences between the two models suggest that problems with model implementation, e.g. input data identification and/or data loss, are more substantial than problems with information extraction.

Model residuals as a function of data loss further indicate that the performance of ANNs is insensitive to the treatment applied to the training data. While the standardized method showed slight improvement as a function of data loss, especially at intermediate levels of loss (Fig. 6), the range of error values under both treatments were similarly large and the standard error estimates from both methods fully enveloped those of the other model. Model performance cannot be discriminated due to the large standard deviations of error based on a diurnal pattern. Since both models yield similar error values for the full time series and similarly over/under-estimate the ET, model improvement from data treatment may be limited.

Our findings are contrary to our initial hypothesis, that altering the data distribution would improve model performance by making more information available during model training. Changing the probability distribution of the data should have improved the sampling rate of extreme events by narrowing the range of values in the standardized index (Fig. 2). Based on the objective functions (Table 2) and the wide variability of errors as a function of data loss (Fig. 6), it appears that the hypothesis may not be true. This result suggests either that the model structure may be flawed or that the missing data presents a much greater obstacle than the transformation can overcome.

4.2 Nocturnal evapotranspiration

One important result when comparing the two models is the difference in nocturnal ET (Table 3). Both models have reduced errors at night, and are similarly prone to over and underestimation, especially when rates of data loss are high (Fig. 6). In a diurnal sense, the standardized model has a tendency to predict higher nighttime ET, especially around dawn, when the model shifts into a mode of increasing ET earlier than the rescaled model.

Several other studies have reported on ET at night from deciduous forest sites in more humid regions (e.g., Novick et al., 2009). Few have reported on nighttime activity in arid riparian systems. At our site, nighttime ET is estimated at 0.06 and 0.09 fraction of daily ET based on the rescaled and standardized models, respectively. These fractions are similar to those reported for humid forests, indicating that semi-arid riparian species with persistent access to groundwater evapotranspire at similar rates as humid forests during the night. Evaporative demand is dramatically reduced at night when radiative forcing is absent and relative humidities rise, even in arid locations.

Comparing nocturnal ET to estimated values of daytime and nighttime evaporation (E) and transpiration (T), we find that nighttime ET is a similar fraction of daytime ET as nighttime T to daytime T (~0.5–0.1, Fisher et al., 2007). This suggests that plant

HESSD
7, 6525–6551, 2010

transpiration scales directly with total ET from day to night. Even under the different model treatments proposed here, the ET_n/ET_d ratios are consistent. As noted in the error analysis, the results of gap filling analysis using two different training data sets does not alter predictions of ET_n/ET_d.

4.3 Near-dawn evapotranspiration

The poor model performance at and near dawn provide strong evidence of the flaws associated with eddy covariance data at those times. Conventional methods of eddy covariance data filtering apply a u^*-filter or other methods based on turbulence theory. However, the greater error at dawn compared to other nighttime intervals, despite similar friction velocities, indicate that the onset of turbulence is problematic for both measurement and modeling. Because the theoretical basis for eddy covariance is built on strong turbulent mixing, measurements during and immediately following periods of high stability pose substantial difficulties which may not be overcome in data-dependent regression models. Further investigation into near-dawn energy and moisture fluxes will provide insight into how ecosystems use water during this critical period.

In the scope of model-data fusion approaches, the question of eddy covariance data quality remains a problem for the research community. Model development and calibration is dependent on continuous data records. However, the poor quality and frequent loss of data, especially at night, may require that model results are taken as the standard for comparison against data. The difficulty is in properly identifying the source of the best information for nighttime fluxes. In this study, we show that altering the way in which the model samples data may not significantly improve model results. We postulate that, in future studies, applying ancillary data streams (e.g. chamber flux measurements or soil moisture data) or using alternate model structures may have a greater effect on gap filling results.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

7, 6525–6551, 2010

Modeling moisture fluxes using artificial neural networks

A. L. Neal et al.

4.4 Implications of this study

Overall, this study points toward several arenas for future study in modeling land-atmosphere interactions and gap filling. Applying principles of information theory can indicate what flux values are the most informative for modeling applications. Gap filling as a modeling exercise attempts to restore information where it is lost. As shown here, increasing the information content of low flux values may not dramatically improve overall model performance, but performance is not dramatically reduced either.

An information theory approach may not satisfy deficiencies inherent to the original data. The poor performance of both models during the near-dawn periods can be linked to the timing of the onset of turbulent mixing as the nocturnal boundary layer breaks up. In this study, both data treatments lead to overestimation of fluxes between 05:30and 08:00 local time, when the stable nighttime air is becoming unstable due to surface warming. The meteorological variables used as model inputs indicate increasing fluxes earlier than the actual data. Because this time period represents the threshold of feasible data collection by eddy covariance, other methods (e.g. leaf-level measurements of transpiration) should be used to corroborate the data or model results.

Because of the differences in turbulent mixing between nighttime and daytime periods, parallel gap filling models may be an appropriate solution to the information extraction problem identified here. Where most approaches use a single model for nighttime and daytime, assuming that the flux mechanisms and controlling variables are consistent throughout, we propose using different models for night and day, focusing on identifying appropriate model structures for each. Applying two models will also provide another avenue to explore filter criteria when stable nighttime air forms.

This study also points out one of the potential problems with model training, that improved model performance under certain conditions (here low flux values) may result in poor performance under other conditions. The "black box" nature of ANN model development means that these tradeoffs in model performance come without the ability

HESSD
7, 6525–6551, 2010

to fully trace model outcomes to changes in the model structure or training. Comparing the diurnal plots for the two models, it appears that the standardized model under-estimated midday ET, which may have been ignored in favor of improving ET performance at other times of the day. The risk of this compensatory effect in model training suggests that different models should be used for filling gaps at night and during the day, especially in light of the different micrometeorological conditions at work in those times. Using separate models would allow one model to track patterns under stable nighttime conditions, while another would follow the flux behavior under turbulent mixing.

Acknowledgements. This research was supported by the Sustainability for Semi-Arid Hydrology and Riparian Areas NSF-Science and Technology Center (SAHRA NSF-STC). Comments from Solomon Hsiang and the reviewers were helpful to improve the manuscript. The authors would like to thank Russ Scott (USDA-ARS) for use of the data as well as his advice and expertise at the research site.

References

Alavi, N., Warland, J. S., and Berg, A. A.: Filling gaps in evapotranspiration measurements for water budget studies: evaluation of a Kalman filtering approach, Agr. Forest Meteorol., 141 (1), 57–66, 2006.

Baldocchi, D., Falge, E., Gu, L. H., Olson, R., et al.: FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, B. Am. Meteorol. Soc., 82(11), 2415–2434, 2001.

Blanken, P. D., Black, T. A., Neumann, H. H., Den, G., Hartog, Yang, P. C., Nesic, Z., Staebler, R., Chen, W., and Novak, M. D.: Turbulent flux measurements above and below the overstory of a boreal aspen forest, Bound.-Lay. Meteorol., 89(1), 109–140, 1998.

Brown-Mitic, C., Shuttleworth, W. J., Harlow, R. C., Petti, J., Burke, E., and Bales, R.: Seasonal water dynamics of a sky island subalpine forest in semi-arid Southwestern US, J. Arid Environ., 69(2), 237–258, 2007.

Dawson, T. E., Burgess, S. S. O., Tu, K. P., Oliveira, R. S., Santiago, L. S., Fisher, J. B., Si-

monin, K. A., and Ambrose, A. R.: Nighttime transpiration in woody plants from contrasting ecosystems, Tree Physiol., 27(4), 561–575, 2007.
Falge, E., Baldocchi, D., Olson, R., Anthoni, P., et al.: Gap filling strategies for long term energy flux data sets, Agr. Forest Meteorol., 107(1), 71–77, 2001.
Fisher, J. B., Baldocchi, D. D., Misson, L., Dawson, T. E., and Goldstein, A. H.: What the towers don't see at night: nocturnal sap flow in trees and shrubs at two AmeriFlux sites in California, Tree Physiol., 27(4), 597–610, 2007.
Gochis, D. J., Brito-Castillo, L., and Shuttleworth, W. J.: Hydroclimatology of the North American monsoon region in Northwest Mexico, J. Hydrol., 316(1–4), 53–70, 2006.
Goulden, M. L., Munger, J. W., Fan, S. M., Daube, B. C., and Wofsy, S. C.: Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy, Glob. Change Biol., 2(3), 169–182, 1996.
Gove, J. H. and Hollinger, D. Y.: Application of a dual unscented Kalman filter for simultaneous state and parameter estimation in problems of surface-atmosphere exchange, J. Geophys. Res.-Atmos., 111, D08S07, doi:10.1029/2005JD00602 2006.
Hui, D. F., Wan, S. Q., Su, B., Katul, G., Monson, R., and Luo, Y. Q.: Gap-filling missing data in eddy covariance measurements using multiple imputation (MI) for annual estimations, Agr. Forest Meteorol., 121(1–2), 93–111, 2004.
Knorr, W. and Kattge, J.: Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling, Glob. Change Biol., 11(8), 1333–1351, 2005.
Kurc, S. A. and Small, E. E.: Soil moisture variations and ecosystem-scale fluxes of water and carbon in semiarid grassland and shrubland, Water Resour. Res., 43, W06416, doi:10.1029/2006WR005011, 2007.
MacKay, D. (Ed.): Information Theory, Inference, and Learning Algorithms, Cambridge University Press, 640 pp., 2003.
Mishra, A. K. and Desai, V. R.: Drought forecasting using feed-forward recursive neural network, Ecol. Model., 198(1–2), 127–138, 2006.
Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., et al.: Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes, Agr. Forest Meteorol., 147(3–4), 209–232, 2007.
Neal, A. L.: Toward a Model-Based Method for Gap Filling Flux Data for a Semi-Arid Site, University of Arizona, Tucson, AZ, 2008.

HESSD
7, 6525–6551, 2010

Modeling moisture fluxes using artificial neural networks

A. L. Neal et al.

Novick, K. A., Oren, R., Stoy, P. C., Siqueira, M. B. S., and Katul, G. G.: Nocturnal evapotranspiration in eddy-covariance records from three co-located ecosystems in the Southeastern US: implications for annual fluxes, Agr. Forest Meteorol., 149(9), 1491–1504, 2009.

Papale, D. and Valentini, A.: A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization, Glob. Change Biol., 9(4), 525–535, 2003.

Pitman, A. J.: The evolution of, and revolution in, land surface schemes designed for climate models, Int. J. Climatol., 23(5), 479–510, 2003.

Richardson, A. D., Hollinger, D. Y., Burba, G. G., Davis, K: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, Agr. Forest Meteorol., 136(1–2), 1–18, 2006.

Scott, R. L., Huxman, T. E., Williams, D. G., and Goodrich, D. C.: Ecohydrological impacts of woody-plant encroachment: seasonal patterns of water and carbon dioxide exchange within a semiarid riparian environment, Glob. Change Biol., 12(2), 311–324, 2006.

Scott, R. L., Edwards, E. A., Shuttleworth, W. J., Huxman, T. E., Watts, C., and Goodrich, D. C.: Interannual and seasonal variation in fluxes of water and carbon dioxide from a riparian woodland ecosystem, Agr. Forest Meteorol., 122(1–2), 65–84, 2004.

Scott, R. L., Cable, W. L., Huxman, T. E., Nagler, P. L., Hernandez, M., and Goodrich, D. C.: Multiyear riparian evapotranspiration and groundwater use for a semiarid watershed, J. Arid Environ., 72(7), 1232–1246, 2008.

Shannon, C. E.: A mathematical theory of communication, Bell Syst. Tech. J., 27, 379–423, 623–656, 1948.

Wohlfahrt, G., Fenstermaker, L. F., and Arnone, J. A.: Large annual net ecosystem CO_2 uptake of a Mojave Desert ecosystem, Glob. Change Biol., 14(7), 1475–1487, 2008.

HESSD

7, 6525–6551, 2010

Modeling moisture fluxes using artificial neural networks

A. L. Neal et al.

Table 1. Shannon information index values of latent heat flux used for training ANN models.

ET (mm)	Hrsc	ET (mm)	Hstd
<10.08	0.2777	<−3.23	0.0009
10.08–38.24	0.1397	(−3.23)–(−1.24)	0.0019
38.24–66.4	0.092	(−1.24)–(−0.1)	0.0022
66.4–94.56	0.0792	−0.1–0.31	0.0758
94.56–122.72	0.0694	0.31–0.39	0.1042
122.72–150.88	0.0644	0.39–0.51	0.0476
150.88–179.04	0.0594	0.51–0.63	0.0287
179.04–207.2	0.0531	0.63–0.75	0.0123
207.2–235.36	0.0511	0.75–1.4	0.1153
235.36–263.52	0.0445	1.4–4.19	0.1172
263.52–291.68	0.0396	4.19–10.75	0.1191
291.68–319.84	0.0293	10.75–22.44	0.1193
319.84–348	0.021	22.44–40.25	0.1998
348–376.16	0.017	40.25–64.37	0.1921
376.16–404.32	0.0138	64.37–94.96	0.1198
404.32–432.48	0.0096	94.96–132.19	0.1094
432.48–460.64	0.007	132.19–176.22	0.1065
460.64–488.8	0.0033	176.22–227.24	0.104
488.8–516.96	0.0035	227.24–285.41	0.0879
516.96–545.12	0.0029	285.41–350.89	0.0554
545.12–573.28	0.0016	350.89–423.86	0.0317
573.28–601.44	0.0007	423.86–504.49	0.0133
601.44–629.6	0.0007	504.49–592.95	0.005
629.6–657.76	0.0002	592.95–689.4	0.0015
>657.76	0.0002	>689.4	0.0002
Total	1.08	Total	1.77

HESSD
7, 6525–6551, 2010

HESSD
7, 6525–6551, 2010

Modeling moisture fluxes using artificial neural networks

A. L. Neal et al.

Table 2. Summary of model performance for both ANNs. Model training used mean absolute error (MAE) to identify parameters.

Metric	Training		Validation	
	ET_{rsc}	ET_{std}	ET_{rsc}	ET_{std}
MAE (mm)	0.0252	0.0329	0.0272	0.0348
RMSE 30 min (mm)	0.0403	0.0517	0.0413	0.0522
RMSEdaily (mm)	0.7395	0.9301	1.0544	1.0876
Rel RMSE (–)	0.0369	0.0475	0.3920	0.4954
Correl	0.8951	0.8326	0.8638	0.7916

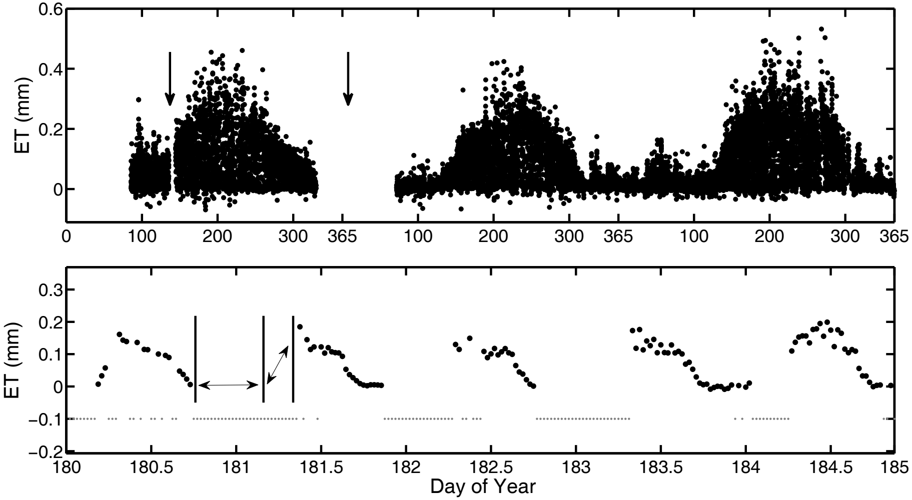

Fig. 1. The time series of latent heat flux used in this study shown as the full record (upper plot) and highlighting in more detail the concerns associated with gap filling, including properly identifying the length of near-zero fluxes at night, and the timing and rate of flux increase during the day (lower plot).

HESSD

7, 6525–6551, 2010

Modeling moisture fluxes using artificial neural networks

A. L. Neal et al.

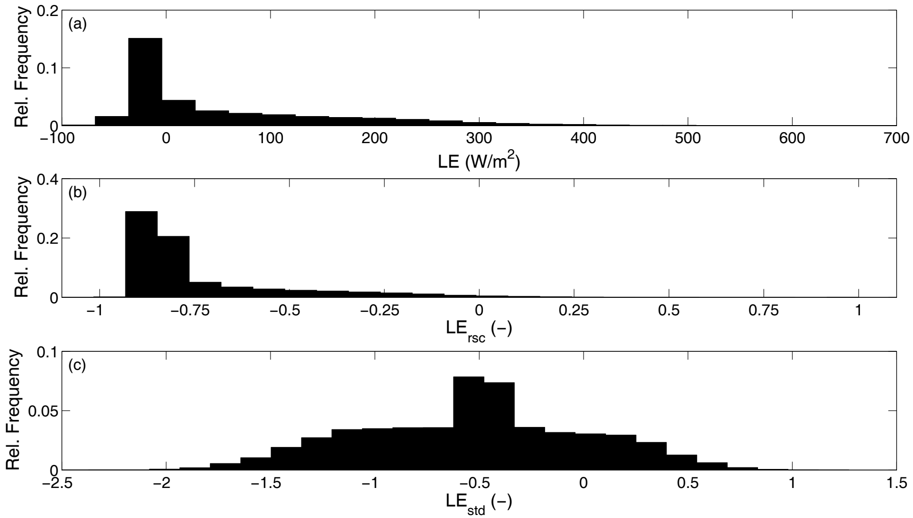

Fig. 2. Histograms of relative frequency of latent heat flux at the flux tower site. Frequency of the raw data is shown in **(a)**. Plot **(b)** is the frequency of the LE_{rsc} (the *mapminmax* function of the Matlab Neural Network Toolbox was used to perform scaling). Plot **(c)** is a histogram of standardized LE as described in Data and Methods. Note that the mapminmax retains the original distribution while the standardization transforms the distribution to something near normal for ET.

HESSD
7, 6525–6551, 2010

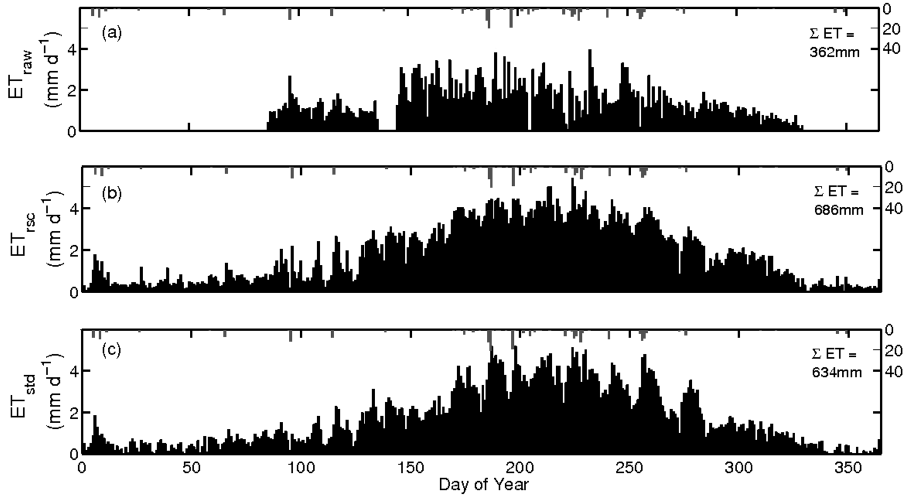

Fig. 3. Precipitation (P) and daily ET for **(a)** ET_{raw}, **(b)** ET_{rsc}, and **(c)** ET_{std} for the year 2001. Sum of ET over the year is shown in the box associated with each plot (sum of precipitation is 263 mm).

HESSD
7, 6525–6551, 2010

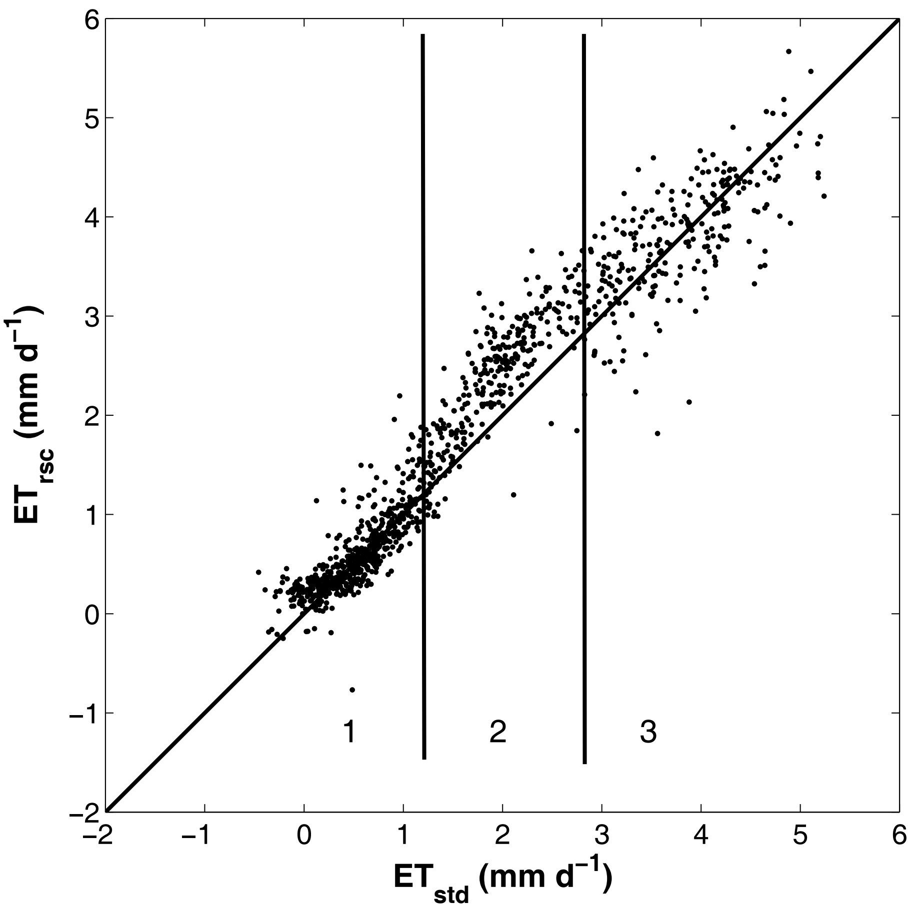

Fig. 4. Scatterplot of ET_{rsc} and ET_{std} with labels indicating good model agreement (1, winter), relative overestimation in ET_{rsc} (2, pre- and post-monsoon), and high intra-model variability (3, monsoon).

HESSD
7, 6525–6551, 2010

Modeling moisture fluxes using artificial neural networks

A. L. Neal et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

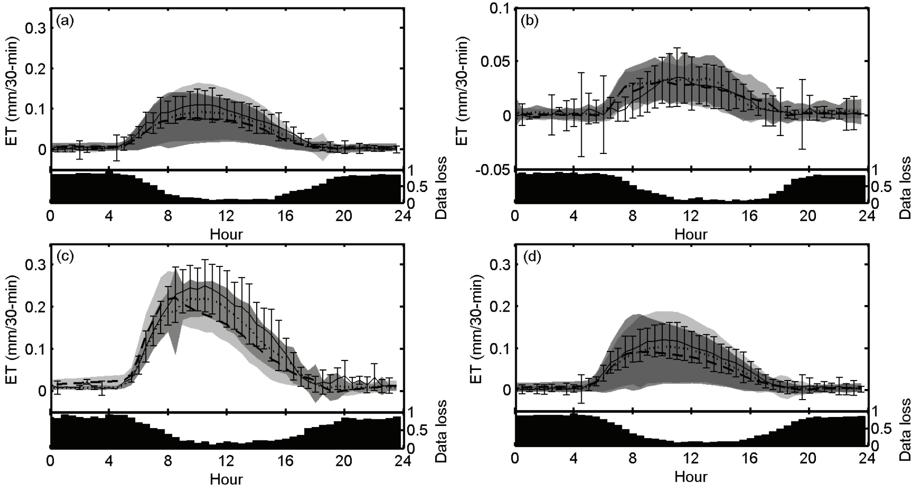

Fig. 5. Diurnal plots from **(a)** pre- and post-monsoon seasons, **(b)** winter, **(c)** monsoon, and **(d)** the full time series. For each plot the upper figure shows the two model results and the raw data. Lower plots show the rate of data loss during that period.

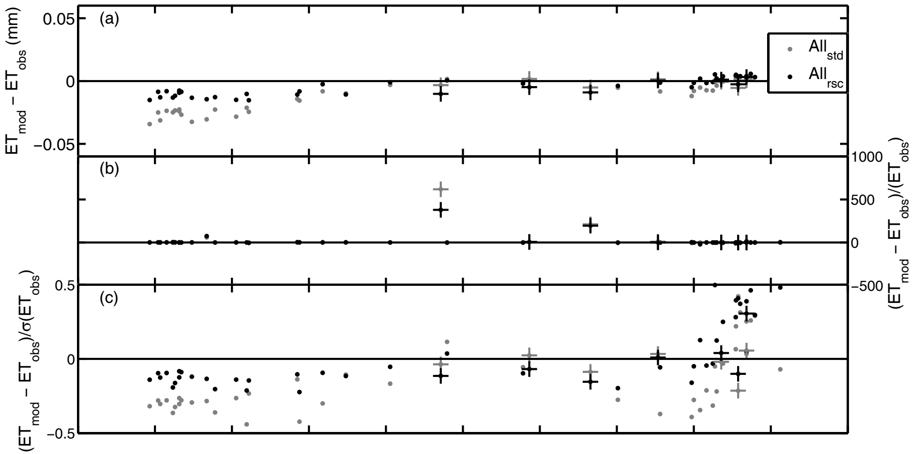

Fig. 6. **(a)** Model residuals, **(b)** normalized model residuals and **(c)** standardized model residuals plotted as a function of data loss for the full data record. Crosses indicate data from 05:30 to 08:00 LT.

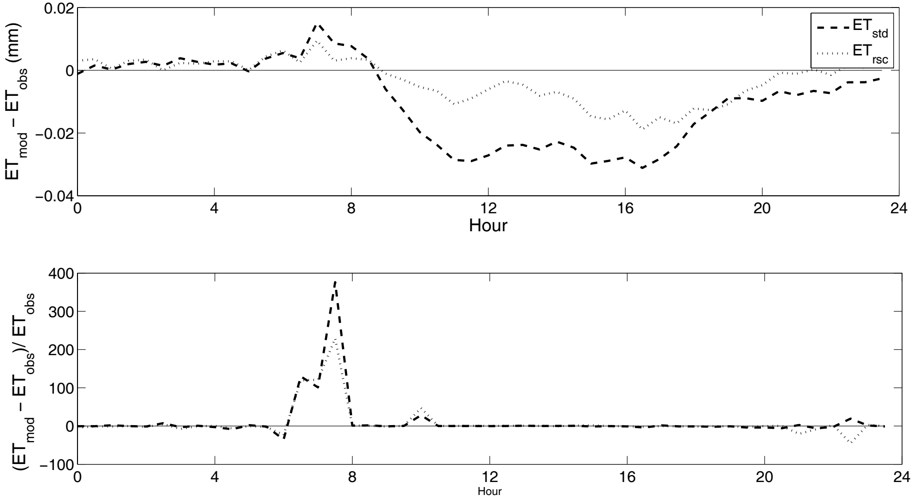

Fig. 7. Diurnal plots of model residuals from the full time series.

HESSD

7, 6525–6551, 2010

Modeling moisture fluxes using artificial neural networks

A. L. Neal et al.